

Description

METHOD AND STRUCTURE FOR CONTROLLING STRESS IN A TRANSISTOR CHANNEL

BACKGROUND OF INVENTION

[0001] The invention generally relates to methods for manufacturing a semiconductor device with improved device performance, and more particularly to methods for manufacturing semiconductor devices which impose tensile and compressive stresses in the substrate of the device during device fabrication.

[0002] Generally, metal-oxide semiconductor transistors include a substrate made of a semiconductor material, such as silicon. The transistors typically include a source region, a channel region and a drain region within the substrate. The channel region is located between the source and the drain regions. A gate stack, which usually includes a conductive material, a gate oxide layer and sidewall spacers, is generally provided above the channel region. More par-

ticularly, the gate oxide layer is typically provided on the substrate over the channel region, while the gate conductor is usually provided above the gate oxide layer. The sidewall spacers help protect the sidewalls of the gate conductor.

[0003] It is known that the amount of current flowing through a channel which has a given electric field across it, is generally directly proportional to the mobility of the carriers in the channel. Thus, by increasing the mobility of the carriers in the channel, the operation speed of the transistor can be increased.

[0004] It is further known that mechanical stresses within a semiconductor device substrate can modulate device performance by, for example, increasing the mobility of the carriers in the semiconductor device. That is, stresses within a semiconductor device are known to enhance semiconductor device characteristics. Thus, to improve the characteristics of a semiconductor device, tensile and/or compressive stresses are created in the channel of the n-type devices (e.g., NFETs) and/or p-type devices (e.g., PFETs). However, the same stress component, for example tensile stress or compressive stress, improves the device characteristics of one type of device (i.e., n-type device or

p-type device) while discriminatively affecting the characteristics of the other type device.

[0005] In order to maximize the performance of both NFETs and PFETs within integrated circuit (IC) devices, the stress components should be engineered and applied differently for NFETs and PFETs. That is, because the type of stress which is beneficial for the performance of an NFET is generally disadvantageous for the performance of the PFET. More particularly, when a device is in tension (in the direction of current flow in a planar device), the performance characteristics of the NFET are enhanced while the performance characteristics of the PFET are diminished. To selectively create tensile stress in an NFET and compressive stress in a PFET, distinctive processes and different combinations of materials are used.

[0006] For example, a trench isolation structure has been proposed for forming the appropriate stresses in the NFETs and PFETs, respectively. When this method is used, the isolation region for the NFET device contains a first isolation material which applies a first type of mechanical stress on the NFET device in a longitudinal direction (parallel to the direction of current flow) and in a transverse direction (perpendicular to the direction of current

flow). Further, a first isolation region and a second isolation region are provided for the PFET and each of the isolation regions of the PFET device applies a unique mechanical stress on the PFET device in the transverse and longitudinal directions.

SUMMARY OF INVENTION

[0007] In a first aspect of the invention, the invention provides a method for manufacturing a device including an n-type device and a p-type device. The method involves forming a shallow-trench-isolation oxide (STI) isolating the n-type device from the p-type device and adjusting the shallow-trench-isolation oxide corresponding to at least one of the n-type device and the p-type device such that a thickness of the shallow-trench-isolation oxide adjacent to the n-type device is different from a thickness of the shallow-trench-isolation oxide adjacent to the p-type device. The method further involves forming a strain layer over the semiconductor substrate.

[0008] In another aspect of the invention, the invention provides a method for manufacturing a device including an n-type device and a p-type device. The method involves forming a boundary for the n-type device and the p-type device and adjusting a height of the boundary such that the

boundary adjacent to the n-type device is at a level which is different from a level of the boundary adjacent to the p-type device. The method further involves forming a strain layer over the semiconductor substrate.

[0009] In another aspect of the invention, the invention provides a semiconductor device having a strain layer formed over n-type transistors and p-type transistors formed on a silicon substrate. The device further includes a shallow-trench-isolation oxide around each of the n-type transistors and the p-type transistors, wherein an upper surface of the shallow-trench-isolation oxide of the n-type transistors is at a level than an upper surface of the shallow-trench-isolation oxide of the p-type transistors.

BRIEF DESCRIPTION OF DRAWINGS

[0010] Figure 1 illustrates desired stress states for PFETs and NFETs;

[0011] Figures 2(a) through 2(l) illustrate an exemplary process for forming a transistor according to the invention;

[0012] Figure 3 illustrates an exemplary transistor having an STI oxide which is substantially planar with the silicon substrate;

[0013] Figure 4 illustrates an exemplary transistor having an STI oxide which has an upper surface which is lower than the

upper surface of the silicon substrate;

[0014] Figure 5 illustrates an exemplary transistor having an STI oxide which has an upper surface which is higher than the upper surface of the silicon substrate;

[0015] Figure 6 is an exemplary simulation showing the stress in the transistor channel depending on the height of the STI oxide;

[0016] Figure 7 is another exemplary simulation showing the stress in the transistor channel depending on the height of the STI oxide; and;

[0017] Figure 8 is an exemplary simulation showing the stress level in the transistor channel when the STI oxide is far from the transistor gate.

DETAILED DESCRIPTION

[0018] The invention provides a method for fabricating devices with improved performance characteristics. Tensile and/or compressive stresses can be provided in the channel of a transistor depending on the proximity of the grown stress/strain layer to the channel of the transistor. In an aspect of this invention, by selectively adjusting the topography of the shallow-trench-isolation oxide, which generally surrounds the transistor, tensile stresses can be provided in the channel of NFETs and compressive

stresses can be provided in the channel of PFETs such that the performance of both the NFETs and the PFETs can be improved using the same stress/strain layer.

[0019] In this invention, a stress layer, such as a SiGe layer, for example, is used to form stresses in a channel of the semiconductor device. When a SiGe layer is grown on a semiconductor layer the surrounding semiconductor material is subjected to tensile stress while the grown SiGe layer is subjected to compressive stress. In particular, a portion of the semiconductor device is put under tensile stress and the SiGe layer is subjected to compressive stress because the SiGe layer has a different lattice structure than the silicon layer. Further, the stress levels resulting from the SiGe stress layer are relatively high (on the order of 1–2GPa).

[0020] However, as discussed above, tensile stresses in the channel area are beneficial to the NFET drive currents while compressive stresses in the channel area are beneficial to the PFET drive currents. In particular, tensile stresses significantly hinder the PFET drive currents. In this invention, the stresses in the PFET are made to be compressive stresses rather than tensile stresses in order to improve the performance of the PFET. Thus, this invention provides

a method for preventing the formation of tensile stresses and/or providing longitudinal compressive stresses along the channel of the PFET while providing tensile stresses along the channel of the NFET to improve the performance of the devices.

[0021] Figure 1 illustrates desired stress states for improving the performance of PFETs and NFETs (see Wang et al., IEEE Tran. Electron Dev., v.50, p.529, 2003). In Figure 1, an NFET and a PFET are shown to have a source region, a gate region and a drain region. The NFET and PFET are shown to have arrows extending outward from the active area to illustrate tensile stresses. The arrows extending inward toward the PFET device are illustrative of compressive forces. More specifically, the outwardly extending arrows, shown extending from the NFET, illustrate a tensile stress that is desired in the transverse and longitudinal directions of the device. On the other hand, the inwardly extending arrows, shown with relation to the PFET, illustrate a desired longitudinal compressive stress.

[0022] The range of stresses needed to influence device drive currents is of the order of a few hundred MPa to a few GPa. The width and the length of the active area of each device are represented by "W" and "L", respectively. It

should be understood that each of the longitudinal or transverse stress components could be individually tailored to provide the performance enhancements for both devices (i.e., the NFET and the PFET).

[0023] Figures 2(a) through 2(l) depict an exemplary process for forming n-type devices according to this invention. As shown in Figure 2(a), a patterned photo-resist layer 205 is deposited over a silicon substrate 200 and the exposed portion of the silicon substrate 200 is doped, for example, with Ge, As, B, In or Sb. For example, the doping concentration of Ge may be, for example, about 1×10^{14} Ge/cm² to about 1×10^{16} Ge/cm². A doped region 207 is formed in the semiconductor substrate 200.

[0024] Then, as shown in Figure 2(a), the patterned photo-resist layer 205 is removed and a mask 210, made of nitride, for example, is deposited on the surface of the semiconductor substrate 200. The mask 210 protects the semiconductor substrate beneath it from being etched during reactive ion etching (RIE). Generally, the mask 210 exposes portions of the semiconductor substrate where shallow trenches are to be formed via RIE.

[0025] As shown in Figure 2(b), RIE is performed to form grooves/trenches 215 in the semiconductor substrate

200. In particular, the location of the formed grooves/ trenches at least partially overlaps a portion of the doped semiconductor region 207 such that when the grooves/ trenches 215 are formed, the doped semiconductor substrate region is exposed. Oxide material is then deposited to fill the trenches, such that devices adjacent to each other on the semiconductor substrate 200 are electrically isolated from each other.

[0026] As shown in Figure 2(c), oxide material 219 is then deposited to fill the trenches and electrically isolate the device from any adjacent device. After filling the trenches with the oxide material, the mask 210 is removed using any known appropriate method, leaving the oxide material 219 to protrude above the surface of the semiconductor substrate 200. Accordingly, the amount of oxide material 219 protruding above the surface of the semiconductor material 200 may be controlled by controlling the thickness of the mask 210. After the mask 210 is removed, as shown in Figure 2(d), chemical mechanical polishing (CMP) is performed to substantially planarize the upper surface 231 of the semiconductor substrate 200.

[0027] Next, as shown in Figure 2(e), the semiconductor device is further fabricated using known methods. For example, as

shown in Figure 2(e), a gate oxide layer 222 is grown on the upper surface 231 of the semiconductor substrate 200. A gate oxide layer 222 of about 10 Angstroms to about 100Å is generally grown. On the gate oxide layer 222, a polysilicon layer 224 is generally deposited using chemical vapor deposition (CVD) to a thickness of about 500 Angstroms to about 1500 Angstroms to form the gate electrode 232 (see Figure 2(f)).

[0028] Patterned photoresist layers (not shown) are then used to define the gate electrodes. A thin layer of oxide (not shown) is then grown on the remaining polysilicon. Patterned photoresist layers (not shown), which are later removed, are used to successively tip (and halo counter dope) implant the n-type and p-type transistors. For n-type transistors, a very shallow and low dose implant of arsenic ions, for example, may be used to form the p-tip (while a Boron implant, for example, may be used for halos). For p-type transistors, a very shallow and low dose implant of BF_2 ions, for example, may be used to form n-tip (while an arsenic implant may, for example, be used for halos).

[0029] Next, as shown in Figure 2(g), the source and drain regions 240 and 241 of the transistor are formed by sub-

jecting the respective portions of the substrate to ion implantation, for example. Patterned photoresist layers (not shown), which are removed prior to the next stage of the process, are used to successively create the source/drain regions of the transistors. For the n-type transistors, a shallow and high-dose of arsenic ions, for example, may be used to form the source/drain regions 240 and 241 while the p-type transistors are covered with the corresponding photoresist layer. As discussed above, in the methods according to this invention, the source and drain regions 240 and 241 are formed in upper portions of semiconductor substrate 200 (i.e., not removed and re-formed).

[0030] For the p-type transistors, a shallow and high dose of BF_2 ions, for example, may be used to form the source/drain regions 240 and 241 while the n-type transistors are covered with the corresponding photoresist layer. An anneal is then used to activate the implants. The exposed oxide on the structure is then stripped by dipping the structure in HF (hydrofluoric acid) in order to expose bare silicon in the source, gate and drain regions of the transistors.

[0031] Further, as shown in Figure 2(h), spacers 228 may be formed by depositing a silicon nitride layer (not shown)

using CVD to a thickness of about 100 Angstroms to about 1000 Angstroms and then etching the nitride from the regions other than the sidewalls of the gate. The combination of the gate oxide layer 222, gate electrode 224 and spacers 228 may be referred to as a gate stack.

[0032] Referring to Figure 2(i), metal 232 is deposited to a thickness of about 30 Angstroms to about 200 Angstroms across the wafer surface in order to form silicide 243. The silicide could be formed from reacting the underlying with any deposited metal such as Co, Hf, Mo, Ni, Pd₂, Pt, Ta, Ti, W, and Zr. In the regions, such as, the source, drain and gate regions, where the deposited metal is in contact with silicon, the deposited metal reacts with the silicon to form silicide.

[0033] Next, the structure is heated to temperature of about 300°C to about 1000°C to allow the deposited silicide material to react with the exposed polysilicon or silicon. During sintering, silicide only forms in the regions where metal is in direct contact with silicon or polysilicon. In the other regions (i.e., where the deposited metal is not in contact with silicon), the deposited metal remains unchanged. This process aligns the silicide to the exposed silicon and is called "self-aligned silicide" or salicide. The

unreacted metal is then removed using a wet etch while the formed silicide remains.

[0034] After salicidation is completed, in the methods according to this invention an etch stop nitride film 245 is deposited over the wafer, as shown in Figure 2(k). The film may be of tensile stress or compressive stress. Depending on the topography of the shallow-trench-isolation oxide (STI) which surrounds and/or isolates one device from another device, the stress level induced in the transistor channel may be changed. Thus, by adjusting the thickness of the shallow-trench-isolation oxide which is adjacent to and/or isolates one device from the others, it is possible to adjust the stress level in the corresponding transistor channel such that the desired stress type (i.e., compressive or tensile) is formed therein.

[0035] As shown in Figure 2(l), after formation of the etch stop nitride layer an insulating layer 250 is formed thereon. The insulating layer may be, for example, boron-phospho silicate glass (BPSG). The insulating layer not only insulates one layer from another layer, as desired, the insulating layer may also serve be provided to planarize the surface of the device. The fabrication process continues as necessary according to the design specifications.

[0036] In the methods according to this invention because the source and drain regions of the semiconductor device are formed on portions of the semiconductor substrate which are undisturbed (i.e., not etched and re-formed), the surface is more favorable to cobalt silicide formation as cobalt silicide. Further, in the methods according to this invention, the thickness of the STI oxide of a transistor is be adjusted and/or the distance of the STI oxide from the channel of the transistor is adjusted so as to control the amount and type of strain (i.e., compressive strain or tensile strain) created in the channel of the transistor. In particular, as further discussed below, by providing an STI oxide with a first thickness around the NFETs and providing an STI oxide with a second thickness around the PFETs, it is possible to prevent the formation of the "undesired" (i.e., tensile stress in the PFET and compressive stress in the NFET) and even form the desired stress (i.e., compressive stress in the PFET and tensile stress in the NFET) with a same stress/strain layer for both the NFETs and the PFETs.

[0037] Figure 3 shows a cross-section of a transistor with a STI oxide which is substantially planar with the silicon substrate. Figure 4 shows a cross-section of a transistor with

a STI oxide which has a thickness which results in the upper surface of the STI oxide being lower than the upper surface of the silicon substrate. Figure 5 shows a cross-section of a transistor with a STI oxide which has a thickness which results in the upper surface of the STI oxide being higher than the upper surface of the silicon substrate. Generally, when the upper surface of the STI oxide is lower than or higher than the upper surface of the silicon substrate, the STI oxide may be up to about 1000 Angstroms lower or higher than the upper surface of the silicon substrate.

[0038] For example, depending on the type strain layer (i.e., tensile or compressive), either the NFETs or the PFETs may have such a structure while the other of the NFETs or the PFETs has an STI oxide which is lower than or higher than the silicon substrate. In another embodiment, one of the NFETs or the PFETs may have an STI oxide which is lower than the silicon substrate and the other of the NFETs or the PFETs may have an STI oxide which is higher than the silicon substrate. As shown in Figure 6, with the same strain layer, the stress in the channel at about 100 Angstroms below the silicon surface varies with the thickness of the STI (i.e., whether the upper surface of the

strain layer is lower than, substantially planar with or higher than the upper surface of the silicon substrate). In the simulation shown in Figure 6, the stress in the transistor is shown when the upper surface of the STI oxide is 500 Angstroms below the upper surface of the silicon substrate, when the upper surface of the STI oxide is substantially planar with the upper surface of the silicon substrate, and when the upper surface of the STI oxide is 500 Angstroms higher than the upper surface of the silicon substrate. As can be seen from Figure 6, which shows the effects when a tensile strain layer is used, when the upper surface of the STI oxide is 500 Angstroms below the upper surface of the silicon substrate, the tensile stress in the channel is higher than when the upper surface of the STI oxide is 500 Angstroms higher than the upper surface of the STI oxide.

[0039] Figure 7 shows a simulation showing compressive stress. In this case, when the upper surface of the STI oxide is 500 Angstroms higher than the upper surface of the silicon substrate, the compressive stress is reduced. In the simulation shown in Figure 7, the STI oxide is about 1100 Angstroms from the gate of the transistor.

[0040] As shown above, generally, the stress in the channel is re-

duced when the upper surface of the STI oxide is higher than the upper surface of the silicon substrate. For this reason, when a tensile strain layer is used, the STI oxide of the PFET will generally have an STI oxide with a thickness which is greater than the thickness of the STI oxide of the NFET. However, the STI oxide of the PFETs and the NFETs may also have substantially the same thickness. In that case, depending on the desired amount of stress in the respective transistor channels, the STI oxide of the PFETs and the NFETs are generally formed such that the upper surface of the STI oxide of the PFETs are higher than the upper surface silicon substrate while the upper surface of the STI oxide of the NFETs are substantially planar to or lower than the upper surface of the silicon substrate.

[0041] Further, as discussed above, the stress formed in the transistor channel is dependent on the distance of the STI oxide from the corresponding transistor channel. As shown in Figure 8, the stress level in the transistor may not be formed and/or affected when the STI oxide is far from the gate. Generally, the STI oxide is about 1500 Angstroms or less from the gate and, preferably, about 1100 Angstroms or less from the gate. As shown in Figure

8, when the STI oxide is far from the gate, the stress level in the transistor channel is substantially the same regardless of whether the upper surface of the STI oxide is higher than, lower than or substantially planar with the upper surface of the silicon substrate.

[0042] As discussed above, in one embodiment of the invention, the upper surface of the STI oxide of one of the NFETs or the PFETs may be higher than the upper surface of the silicon substrate and the upper surface of the other of the NFETs or the PFETs may be below the upper surface of the silicon substrate. Thus, the thickness of the higher STI oxide may be about 300 Angstroms to about 1000 Angstroms greater than the thickness of the lower STI oxide.

[0043] The thickness of the strain layer may also contribute to the amount of stress/strain generated in the channel of the transistor. Generally, the stress/strain layer has a thickness of about 250 Angstroms to about 1500 Angstroms.

[0044] As discussed above with regards to Figure 1, in PFETs, a longitudinal compressive stress is desired. The typical range for the desired compressive/tensile stresses is on the order of a few hundred MPa to a few GPa. For exam-

ple, stresses of about 100MPa to about 2 or 3 GPa are generally desired. The invention can produce very high compressive stresses and tensile stresses in the channels of the PFET and NFET devices, respectively.

[0045] By providing tensile stresses to the channel of the NFET and compressive stresses to the channel of the PFET the charge mobility along the channels of each device is enhanced. Thus, as described above, the invention provides a method for providing compressive stresses along the longitudinal direction of the channel of a PFET while providing tensile stresses in the channel of an NFET by providing, for example, a blanket strain layer on the semiconductor device. This invention also provides a method for optimizing the stress level in the transistor channel by adjusting the thickness of the shallow-trench-isolation oxide and/or adjusting the location of the shallow-trench-isolation oxide with respect to the channel of the corresponding transistor.

[0046] While the invention has been described in terms of embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims.